Proteins with large influence on network dynamics evolve slowly

Ryan N. Gutenkunst

Center for Nonlinear Studies, Los Alamos National Laboratory, ryang@lanl.gov

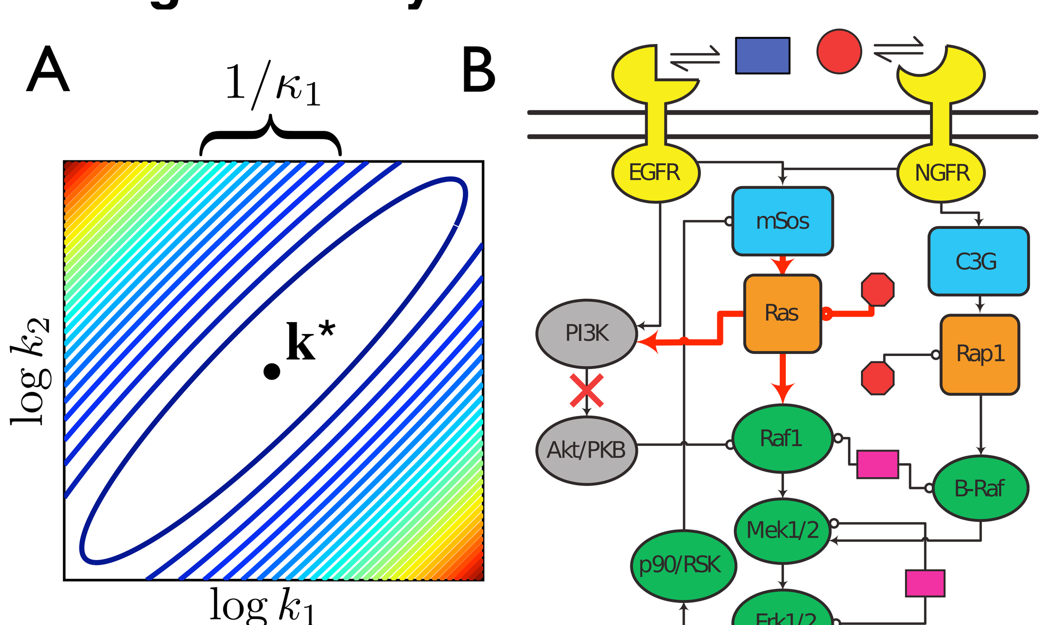
Summary

Studies of protein evolution typically take a coarse view of protein function. Here I take a much finer view, using detailed mechanistic models to measure the influence of each constituent protein on its network's dynamics.

I show that a protein's dynamical influence and its evolutionary rate are negatively correlated, implying purifying selection on network dynamics. Correlation coefficients are typically of order -0.3, among the strongest known correlates of evolutionary rate. Moreover, this correlation is independent of interaction degree, expression level, and knock-out essentiality.

Dynamical influence

Figure 1 - Dynamical influence



The inclusive measure of dynamics, χ^2, sums over all dynamical species in the network [Gutenkunst (2007)].

$$\chi^2(\mathbf{k}) \propto \sum_y \int \left(\frac{y(t,\mathbf{k}) - y(t,\mathbf{k}^*)}{\sigma_y}\right)^2 \mathrm{d}t$$

The dynamical influence κ_i of parameter i is (see **Figure 1A**)

$$\kappa_i = \sqrt{\frac{\partial^2 \chi^2}{\partial^2 \log k_i}}.$$

The dynamical influence $\mathcal{I}$ of a protein is the mean influence of parameters governing reactions the protein participates in.

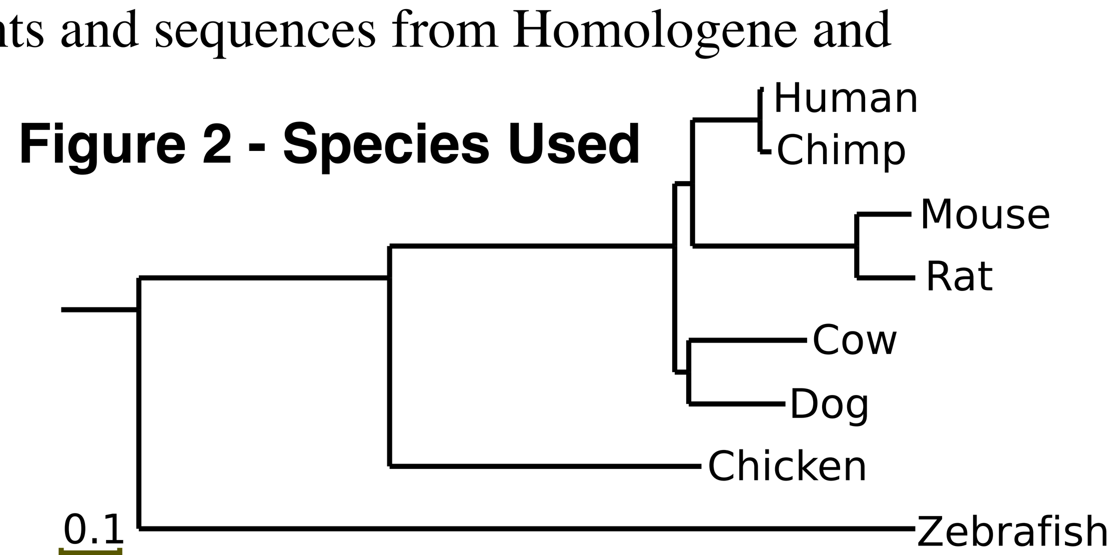
$$\mathcal{I} = \langle \kappa_i \rangle_{\text{geom}}$$

For example, in the network shown in **Figure 1B** the influence of Ras incorporates all the highlighted reactions.

Evolutionary rate

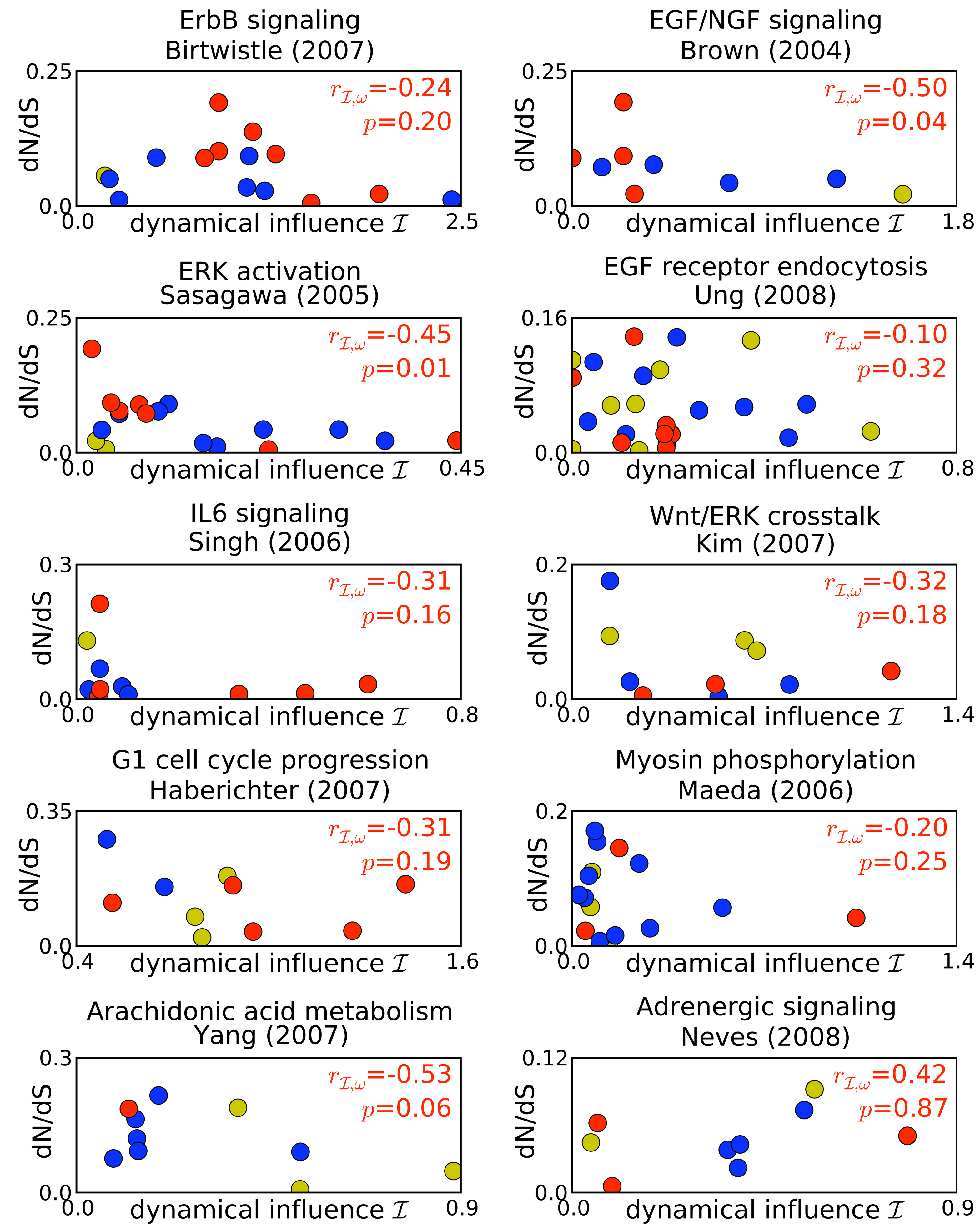
Maximum-likelihood dN/dS values were inferred with PAML, using alignments and sequences from Homologene and GenBank.

Figure 2 - Species Used



Data

Figure 3 - Correlation in real networks



Knock-out essentiality in mouse:
-essential -nonessential -no data

I tested all models in the BioModels database possessing 8 or more species annotated with Uniprot identifiers. Results are shown in **Figure 3**. As seen, for 9 out of 10 models there is a negative correlation between evolutionary rate and dynamical influence. Individual p-values are not dramatic, because model development is inherently low-throughput. However, it is compelling that so many systems show the expected correlation.

Note, the first 5 models all include an experimentally manipulated protein ligand, which I have removed from these analyses. I also omit the influence of multi-protein complexes, as it is not annotated which protein in a complex is active for each reaction. *p*-values are from one-sided permutation tests.

Other potential correlates

It is known that protein evolutionary rates correlate with number of interaction partners (interaction degree), level of expression, and knock-out gene essentiality. I use the number of reactions *R* a protein is involved in (including complexes) as a proxy for interaction degree. I use the total amount *X* of a protein in the model as a proxy for expression. Essentiality in mouse *K* is measured using knock-out data from the Mouse Genome Database. **Table 1** compares the correlations with and without controlling for these variables. It can be seen that these variables typically have little influence on the strength of correlation.

Table 1 - Correlations and partial correlations

Model	$r_{\mathcal{I}\omega}$ (*p*-val)	$r_{\mathcal{I}\omega,R}$	$r_{\mathcal{I}\omega,X}$	$r_{\mathcal{I}\omega,K}$
Birtwistle (2007)	-0.24 (0.20)	0.21	0.14	-0.31
Brown (2004)	-0.51 (0.04)	-0.38	-0.64	-0.56
Sasagawa (2005)	-0.45 (0.01)	-0.45	-0.46	0.21
Ung (2008)	-0.10 (0.32)	-0.15	-0.07	-0.47
Singh (2006)	-0.31 (0.16)	-0.35	-0.48	-0.36
Kim (2007)	-0.32 (0.17)	-0.34	-0.36	-0.25
Haberichter (2007)	-0.31 (0.19)	-0.50	-0.69	-0.10
Maeda (2006)	-0.20 (0.26)	-0.18	-0.22	-0.22
Yang (2007)	-0.53 (0.06)	-0.54	-0.65	-0.15
Neves (2008)	0.42 (0.87)	0.60	0.55	0.35

References

- Birtwistle MR, Hatakeyama M, Yumoto N, Ogunnaike BA, Hoek JB, Kholodenko BN "Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses" *Mol Syst Biol* 3:144 (2007)
- Brown KS, Hill CC, Calero GA, Myers CR, Lee KH, Sethna JP, Cerione RA "The statistical mechanics of complex signaling networks: nerve growth factor signaling" *Phys Biol* 1:184 (2004)
- Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP "Universally sloppy parameter sensitivities in systems biology" *PLoS Comput Biol* 3(10):e189 (2007)
- Haberichter T, Mädge B, Christopher RA, Yoshioka N, Dhiman A, Miller R, Gendelman R, Aksenov SV, Khalil IG, Dowdy SF "A systems biology dynamical model of mammalian G1 cell cycle progression" *Mol Syst Biol* 3:84 (2007)
- Kim D, Rath O, Kolch W, Cho KH "A hidden oncogenic positive feedback loop caused by crosstalk betIen Wnt and ERK Pathways" *Oncogene* 26:4571 (2007)
- Maeda A, Ozaki Y, Sivakumaran S, Akiyama T, Urakubo H, Usami A, Sato M, Kaibuchi K, Kuroda S. "Ca-independent phospholipase A2-dependent sustained Rho-kinase activation exhibits all-or-none response" *Genes Cells* 11:1071 (2006)
- Neves SR, Tsokas P, Sarkar A, Grace EA, Rangamani P, Taubenfeld SM, Alberini CM, Schaff JC, Blitzer RD, Moraru II, Iyengar R "Cell shape and negative links in regulatory motifs together control spatial information flow in signaling networks" *Cell* 133:666 (2008)
- Sasagawa S, Ozaki Y, Fujita K, Kuroda S. "Prediction and validation of the distinct dynamics of transient and sustained ERK activation" *Nat Cell Biol* 7:365 (2005)
- Singh A, Jayaraman A, Hahn J "Modeling regulatory mechanisms in IL-6 signal transduction in hepatocytes" *Biotechnol Bioeng* 95:850 (2006)
- Ung CY, Li H, Ma XH, Jia J, Li BW, Low BC, Chen YZ "Simulation of the regulation of EGFR endocytosis and EGFR-ERK signaling by endophilin-mediated RhoA-EGFR crosstalk" *FEBS Lett* 582:2283 (2008)
- Yang K, Ma W, Liang H, Ouyang Q, Tang C, Lai L. "Dynamic simulations on the arachidonic acid metabolic network" *PLoS Comput Biol* 3:e55 (2007)

Los Alamos NATIONAL LABORATORY EST. 1943

Center for Nonlinear Studies